

Disseksjon av genetikken ved multifaktoriell sykdom – Type 1 diabetes som modell

Kjersti Skjold Rønningen

Divisjon for epidemiologi, Nasjonalt folkehelseinstitutt, Oslo

Korrespondanse: Kjersti Skjold Rønningen, EPSY, Divisjon for epidemiologi, Postboks 4404 Nydalen, 0403 Oslo
Telefon: 22042302 Telefax: 22353605 E-post: kjersti.skjold.ronningen@fhi.no

SAMMENDRAG

Type 1 diabetes skyldes en selektiv autoimmun ødeleggelse av de insulinproduserende cellene i bukspyttkjertelen. Både arvelige og miljømessige faktorer er bestemmende for hvem som rammes av sykdommen. Hva i miljøet som gir startskuddet til ødeleggelse av de insulinproduserende cellene, er fortsatt ukjent. Når det gjelder den arvelige komponenten har det skjedd en rivende utvikling de siste 15 årene. Type 1 diabetes var den første sykdom med multifaktoriell årsakssammenheng det ble utført genom scan for. Genetikken ved type 1 diabetes er kompleks. HLA gener på den korte arm av kromosom nr. 6 er de som bidrar mest, men også mange non-HLA gener er involvert.

Rønningen KS. **Dissecting the genetics of multifactorial diseases – Type 1 diabetes as a model.**

Nor J Epidemiol 2002; 12 (2): 131-135.

ENGLISH SUMMARY

Type 1 diabetes results from a selective autoimmune destruction of the insulin-producing cells of the pancreas. Both genetic and environmental factors are involved in development of the disease. The environmental factors initiating and propagating the process leading to clinical disease are mostly not known. With respect to the genetic component enormous knowledge has been gained during the last 15 years. Type 1 diabetes was the first multifactorial disease where a genome wide screen was performed. The genetics of type 1 diabetes is complex. Most important are the HLA genes on the short arm on chromosome no. 6, but also many non-HLA genes are involved.

FAMILIÆR OPPHOPNING

Høyere konkordans for type 1 diabetes hos eneggede tvillinger (30-50%) sammenliknet med toeggede tvillinger (4,8-27%) og søsken (4,4-12,5%), forteller at risiko for sykdommen delvis er genetisk betinget (1-3). Dette støttes også av at sykdommen viser familiær opphopning. Graden av familiær opphopning (λ_s) estimeres basert på risiko for søsken til en pasient dividert med befolkningsprevalensen. For type 1 diabetes er $\lambda_s = 15$ (6%/0,4%) (4).

IDDM1: HLA GENER

Personer med type 1 diabetes har oftere enn friske spesielle HLA gener. HLA genene koder for våre HLA molekyler, også kalt vevstypemolekyler. HLA systemet, som finnes på den korte arm av kromosom nr. 6 (6p21), er det mest polymorfe gensystem vi kjenner. Bestemte HLA-DQ molekyler bestemmer den viktigste arvelige disposisjonen for type 1 diabetes (5). DQ molekylerne er bygget opp av en variabel alfa kjede (kodet fra et DQA1 gen) og en variabel beta kjede (kodet fra et DQB1 gen). Disposisjonen for å utvikle

type 1 diabetes er bestemt både av DQA1 og DQB1 genene; dvs. hvilke komplette DQ molekyler de koder for. Noen DQ molekyler gir økt risiko for type 1 diabetes, mens andre har beskyttende effekt. DQA1*0102-DQB1*0602 (tidligere kalt DQ6), DQA1*0103-DQB1*0603 (DQ6) og DQA1*0501-DQB1*0301 (DQ7) er assosiert med beskyttelse mot type 1 diabetes, mens DQA1*0301-DQB1*0302 (DQ8) og DQA1*0501-DQB1*0201 (DQ2) er assosiert med økt type 1 diabetes risiko. Den største risikoen for å få sykdommen har personer som er positive både for DQA1*0301-DQB1*0302 og DQA1*0501-DQB1*0201, som har DQA1*0301-DQB1*0302/DQA1*0501-DQB1*0201 genotypen.

For å etablere best mulig risikobestemmelse blant nordmenn, ble 797 tilfeldig valgte pasienter og 469 tilfeldige kontroller typet for alle kjente DQA1 og DQB1 gener. I dette store datasettet fremkom det tydelig at diabetes risiko er bestemt av hvilke DQ genotype du bærer, det vil si DQ genene arvet fra begge foreldre. Som det fremkommer i tabellen under, er det bare seks DQ genotyper som gir statistisk signifikant risiko for type 1 diabetes (6). DQB1*0302/X (X er ikke DQB1*0201, 0302, 0402, 0501, 0604 eller 0602)

ble funnet hos 7,7% av pasientene med type 1 diabetes sammenliknet med 9,2% av kontrollene. Genetisk risiko for type 1 diabetes assosiert med DQB1*0302 og DQB1*0201 i den norske befolkning, utgjøres derfor av de seks DQ genotypene listet i Tabell 1.

Mens spesielle DQ genotyper står for den primære type 1 diabetes assosiasjonen i HLA regionen, modifierer DR4 subtypen som befinner seg på samme kromosom som DQB1*0302 risikoen. DRB1*0401 gir høy diabetes risiko, DRB1*0404 har beskyttende effekt og DRB1*0403 sterk beskyttende effekt. En person som bærer DRB1*0401 istedenfor DRB1*0404 har fem ganger så høy risiko for type 1 diabetes (7).

Tabell 1. Seks ulike DQ genotyper (kombinasjoner arvet fra foreldrene, en DQA1-DQB1 fra mor og en fra far) gir økt risiko for type 1 diabetes. DQA1*0301-DQB1*0302 går igjen i 5 ulike kombinasjoner. For å lette oversikten står DQA1*0301-DQB1*0302 på linjen over hver enkelt av kombinasjonene.

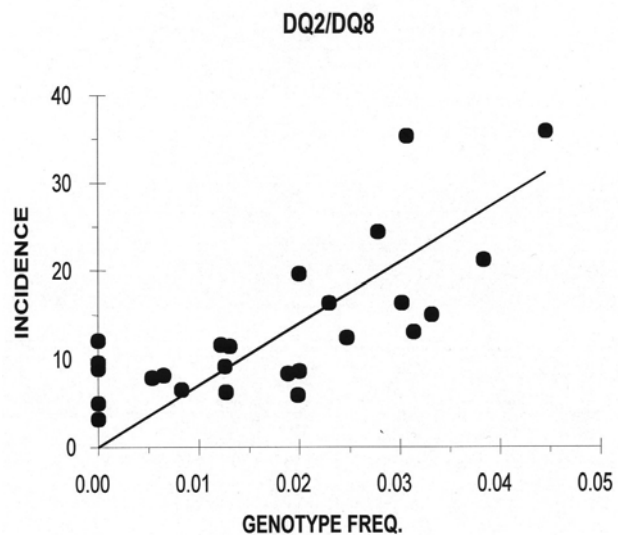
	Relativ risiko	p
DQA1*0301-DQB1*0302		
/DQA1*0501-DQB1*0201	18,6	<0,000001
/DQA1*0401-DQB1*0402	12,7	<0,00001
/DQA1*0301-DQB1*0302	5,8	<0,00001
/DQA1*0101-DQB1*0501	3,1	<0,001
/DQA1*0102-DQB1*0604	2,8	<0,006
DQA1*0501-DQB1*0201		
/DQA1*0501-DQB1*0201	3,2	<0,002

HLA-RISIKOGENER OG INSIDENS

12,4% av nordmenn har en av de seks DQ genotypene som er vist i Tabell 1. Dette er med på å forklare hvorfor Norge er et høy-insidens land for type 1 diabetes. På europeisk basis har vi vist god korrelasjon mellom forekomst av DQA1*0301-DQB1*0302/DQA1*0501-DQB1*0201 (DQ8/2) genotypen i bakgrunnsbefolkningen og insidens av type 1 diabetes i det enkelte land, $p < 0,001$ (8), se fig. 1.

IDDM2: INSULINGENET

Det er også vist at insulingenet er assosiert med type 1 diabetes. Dette genet befinner seg på den korte arm av kromosom nr. 11 (11p15.5), og det er den regulatoriske komponenten VNTR (variable number of tandem repeats), som befinner seg oppstrøms for insulingenet, som er assosiert med type 1 diabetes (9). VNTR er involvert i reguleringen av transkripsjonen av insulingenet. Klasse I variantene (26-63 repeats) er assosiert med sykdomsrisiko, klasse III variantene (140-210 repeats) er assosiert med beskyttelse. Å ha to klasse I varianter gir en relativ risiko for type 1 diabetes på 2,3. Tilstedeværelse av klasse III VNTR gir et høyere transkripsjonsnivå av insulin i tymus, og dette kan fasilitere induksjon av immunologisk toleranse ved å stimulere til negativ seleksjon av insulin-spesifikke T-celler.



Figur 1. Insidens av type 1 diabetes og forekomst av høyrisikogenotypen DQ2/8 (DQA1*0501-DQB1*0201/DQA1*0301-DQB1*0302) i 25 europeiske land.

Insidens: Nye tilfelle per 100 000/år i aldergruppen <15 år. Totalt inngikk 18 417 barn som hadde fått type 1 diabetes.

Genotypedata: Prosent positive for DQ2/8 i hver enkelt befolkning. Totalt inngikk 11 736 DQ typede prøver.

IDDM1 OG IDDM2 – IKKE HELE FORKLARINGEN

I jakten etter gener involvert i utvikling av type 1 diabetes, lette en først etter gener der en forventet effekt, såkalt kandidatgen tilnærming. IDDM1 og IDDM2, som ble funnet på denne måten, er imidlertid ikke nok til å forklare den familiære opphopning av type 1 diabetes. λ_s for et gitt sykdomslocus (gen med alleliske varianter) kan estimeres basert på forholdet mellom forventet andel av affiserte søsken som ikke deler noen alleler ved såkalt "identical-by-descent" (IBD), som er 0,25, og observert andel. For type 1 diabetes multiplex familier fant en at to affiserte søsken ikke delte noe IDDM1 allel (en TNFA mikrosatelitt ble brukt) i 0,8% av tilfellene. λ_s for IDDM1 blir da 0,25 dividert med 0,08 som er lik 3,1 (10). For IDDM2 kunne en i det samme materialet rekne ut λ_s til 1,3. Fra disse lambda-verdiene kan det utledes at IDDM1 og IDDM2 er ansvarlig for henholdsvis 42% og 10% av den familiære opphopning av type 1 diabetes (4), se fig. 2.

HELGENOM SØK ETTER RISIKOGENER

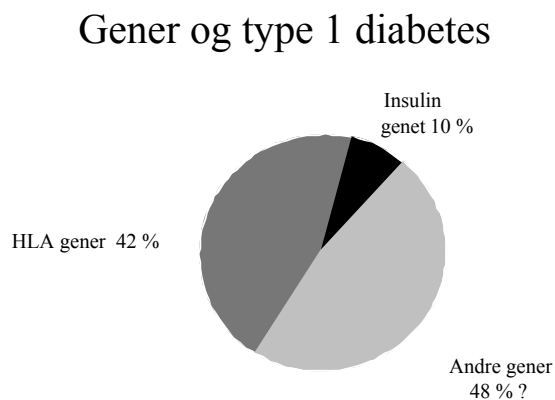
Ved hjelp av mer og mer automatisert teknikk er det karakterisert stadig flere polymorfe markører (først og fremst mikrosatelitter), og samtidig sekvensert stadig større områder av det humane genom. Samtidig har det skjedd en stormende utvikling av computerbaserte nettverkssystemer som gir rask tilgang til informasjon om nye genetiske markører gjennom søk i databaser. Da det også er utviklet fluorescensbasert automatisert DNA fragmentbestemmelsesteknikk, kan man i dag meget en-

kelt og raskt finne polymorfe genetiske markører i nesten alle kromosomale regioner. Dette har gjort genetisk mapping, også av arvemessig komplekse sykdommer som type 1 diabetes mulig.

MULTIPLEX FAMILIER

For at helgenom søk skal kunne identifisere type 1 diabetes disponerende gener, må en ha tilgang til familier der det finnes mer enn ett medlem med sykdommen. Både familier med tilfeller av type 1 diabetes i flere generasjoner, familier med minst to søsken med type 1 diabetes og foreldre uten sykdommen, samt familier som består av minst to affiserte søsken og friske søsken kan brukes. Det mest informative er multiplex familier som består av minst to søsken med type 1 diabetes og to foreldre uten sykdommen. En velger ut mikrosatelitter som dekker hele det humane genom til koblingsanalyser (linkage analyser) (10). Mikrosatelitt-teknikken baserer seg på at det i det humane genom finnes repeterende stereotype DNA-sekvenser (som oftest "dinucleotide repeats", f.eks. (CA)_n; mikrosatelitter). Disse er svært polymorfe med hensyn på lengde og kan derfor brukes som genetiske markører. Ved linkage-analyser ser en på hvorvidt en bestemt genetisk markør nedarves sammen med sykdom. Når familier med to barn med type 1 diabetes studeres, ser en på hvorvidt de to søskenene oftere enn forventet har arvet en mikrosatelitt med lik lengde (såkalt "identity by descent" (IBD)-analyse).

Figur 2. Type 1 diabetes viser familiær opphopning. HLA



gener og insulingenet kan kun forklare 52% av dette. Det må derfor også være andre gener involvert. At så mye som 48% utgjøres av ukjente gener er imidlertid ikke sikkert, da familiær opphopning også kan forklares ved at søsken i stor grad deler miljø.

20 KROMOSOMALE REGIONER

Allerede i 1994 ble det publisert studier der det var utført genom scan for type 1 diabetes gener (10,11). Den sterkeste koblingen ble funnet til IDDM1. Kobling ble også funnet for mikrosatelitter i insulin-gen regionen (IDDM2). 18 andre kromosomale regioner viste holddepunkter for kobling i det første engelske multiplex fami-

lie materialet. Dette gjaldt kromosom 2q (q=lange arm), 3q, 6q (tre ulike regioner), 7q, 8q, 10cen (centromert på kromosomet), 10q, 11q, 13q, 14q, 16q, 17p (p=korte arm), 18q, 19q og X kromosomet (to distinkte regioner). Kromosom 11q og 6q ble videre studert. Den type 1 diabetes assosierte regionen på kromosom 11q ble lokalisert til en 14 centiMorgan (1 cM = 1 million DNA baser) region og fikk navnet IDDM4. På kromosom 6q var det ESR mikrosatelitt markøren som gav sterkest linkage, og denne regionen fikk navnet IDDM5. En kanadisk forskningsgruppe brukte litt andre mikrosatelitter og fant sykdomskobling til en markør på den lange arm av kromosom 15 (15q26, ref. 12). Dette området ble kalt IDDM3.

STORE OMRÅDER

Linkageanalyser (koblingsanalyser) har den store fordel at de ikke forutsetter koblingsulikevekt mellom den aktuelle genetiske markør og det sykdomsdisponerende gen; dvs. man behøver ikke å forutsette at bestemte varianter av markøren befinner seg koblet til bestemte alleler av det sykdomsdisponerende gen. Ulempen med metoden er lav sensitivitet og at man ender opp med store genetiske områder. Dette betyr at når en har funnet linkage mellom en mikrosatelitt og type 1 diabetes, kan man kun si at i et område på et bestemt kromosom (som ofte vil strekke seg over flere millioner basepar DNA) ligger det et gen som har betydning for utvikling av type 1 diabetes. En finner hverken hvilket gen det er eller hvilken betydning markørregionen har for utvikling av type 1 diabetes.

Mens linkage mellom et markørlokus (region) og type 1 diabetes betyr at det aktuelle genet befinner seg innenfor et område på 10-20 centiMorgans, vil påvisning av assosiasjon mellom en markør og type 1 diabetes definere genets lokalisasjon mye nærmere; dvs. genet befinner seg innenfor et område på 2 centiMorgan.

NÆRMERE LOKALISERING – ASSOSIASJONS-STUDIER I SIMPLEX FAMILIER

I den første genom scan studien ble det funnet holddepunkt for kobling til en region på kromosom 2's lange arm (10). Ved hjelp av assosiasjonsstudier, transmisjonsskjevfordelingstesting (transmission disequilibrium test = TDT tesing) og bruk av flere mikrosatelitter innenfor det området det var funnet kobling ble IDDM7 identifisert (13). IDDM7 befinner seg ikke lenger ifra en mikrosatelitt som kalles D2S152 enn 2 centiMorgans. D2S152 er lokalisert i en slik posisjon på kromosom 2q (beregnet til 2q31) at det tyder på at IDDM7 faktisk kan ligge enda nærmere D2S152, kanskje bare 300 kb (0,3 centiMorgans) fra denne mikrosatelitten. Meget interessant i denne sammenheng er at genet som koder for type 1 diabetes autoantigenet glutaminsyre dekarboksylase 67 (GAD1) befinner seg i nettopp denne regionen. Til å lokalisere IDDM6 til 18q12-q21, ble det benyttet 1067 familier fra 4 forskjellige land. Inkludert i disse studiene

var 380 norske simplex familier (et barn med type 1 diabetes, foreldre og friske søsken, ref. 14). Samme strategi ble brukt til finlokalisering av IDDM10. Her befinner genet seg innenfor et område på 1 million basepar (15).

ANNEN GENERASJONS SCAN

Etter dette har det blitt publisert et stort antall studier der det har blitt utført genom scan på multiplex familiematerialer. IDDM8, IDDM9, IDDM11, IDDM12, IDDM13 og IDDM15 er nå også identifisert. I tillegg er det funnet linkage for et område på kromosom 14 og et på kromosom 16 som ikke har fått noe IDDM nummer (referanser i (16)). For mange av "IDDMene" har også linkage blitt konfirmert i andre studier. Et problem har imidlertid vært mangel på nye multiplex familier. Linkage har derfor ved flere anledninger vært repetert i et delvis overlappende familiemateriale.

VALG AV SIGNIFIKANSNIVÅ

Ved random screening av det humane genom for polygene sykdommer, er valg av signifikansnivå for linkage kritisk. Tradisjonelt har linkage studier krevd en lod score >3 for å bli betraktet som statistisk signifikante. Ved slik strategi fant en for type 1 diabetes bare signifikant lod score for HLA. Nitten andre områder hadde lod score >1 . Setter en signifikansnivået ved lod score >1 , øker sensitiviteten, men det er stor sjanse for falsk positivitet. Videre støtte for initiale linkage funn kan imidlertid finnes ved å undersøke andre uavhengige datasett. En stor styrke i denne sammenheng, er tilgang til store datasett av multiplex familier fra etnisk homogene populasjoner.

SKANDINAVISK EU-STØTTET SAMARBEIDSTUDIE

Tilgang på etnisk homogene populasjoner er grunnlaget for samarbeidsstudien "Genome-wide characterization of genetic susceptibility to insulin-dependent diabetes mellitus (IDDM) in man". I denne studien har Danmark, Sverige og Norge samarbeidet om å skaffe et størst mulig antall multiplex familier. Totalt lyktes det å samle 424 multiplex familier, derav 408 med minst to affiserte søsken. 98 norske familier med DNA prøver fra 487 personer har blitt analysert. Også friske søsken har vært inkludert. Det ble genotypet for 324 markører (17).

Som i andre "genom scan" studier ble det også i våre analyser brukt tilleggsfamilier for teste for heterogenitet i avstand mellom genetiske markører sammenliknet med

andre familiematerialer. 126 familier med 158 affiserte søskenpar var tilgjengelige fra Frankrike, og 255 familier med 310 affiserte søskenpar var tilgjengelige fra USA.

HVA FANT VI I SKANDINAVIA?

I tillegg til at vi har verifisert IDDM1 og IDDM2 som type 1 diabetes gener, fant vi høy signifikant kobling til IDDM15 på kromosom 6q. Bortsett fra IDDM15, fant vi ikke non-HLA type 1 diabetes disponerende loci med lod score over 0,8. Det ble imidlertid funnet holdepunkt for tilleggs-loci på kromosomene 2p, 5q og 16p. Regionen på den korte arm av kromosom 2 ligger nær genet som er mutert ved Wolcott-Rallison syndrom (neonatal insulin-avhengig diabetes og epifysal dysplasi). Regionen er klart en annen enn de tidligere identifiserte genområder på kromosom 2 (IDDM7, IDDM12 og IDDM13). Regionen på 5q har ikke tidligere funnet koblet til type 1 diabetes, genområdet er ikke IDDM18. Regionen er imidlertid nylig funnet koblet til multipel sklerose. På den korte arm av kromosom 16 fant vi linkage til to ulike områder. Ikke for noen av disse markørene har det tidligere vært vist linkage til type 1 diabetes. Det ble funnet holdepunkter for interaksjon mellom både IDDM1 og IDDM2 og hvert av områdene på kromosom 16p.

VEIEN VIDERE

I undersøkelser der det blir utført helgenom koblingsundersøkelser av multiplex familier finner alle at det er sikker og sterk kobling til IDDM1 og IDDM2. For å verifisere linkage til de 18 andre regionene som kan være involvert, er det avgjørende med etnisk homogene og store datasett. Av stor betydning er det også med kritisk dataanalyse og bestemmelse av signifikansnivå. I vår nylig avsluttede undersøkelse av vårt store antall familier fra Norge, Sverige og Danmark fant vi kun signifikant linkage til IDDM15. For å komme videre i jakten på nye type 1 diabetes disponerende gener, trengs enda flere multiplex familier.

EUROPEAN CONSORTIUM FOR IDDM GENOME STUDIES

Gruppen består av følgende medlemmer (medlemmer av styringsgruppen er merket med stjerne): Pernilla Holm, Cecilie Julier*, Ingrid Kockum, Valerie Senee, Jenny Papp, Karin Akeson, Christos Bartsocas*, Alberto de Leiva*, Gisela Dalquist*, Kjersti S. Rønningen*, Mark Lathrop, Holger Lutman*, Flemming Pociot* og Jørn Nerup*.

REFERANSER

1. Kaprio J, Tumiletho J, Koskenvuo M, et al. Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. *Diabetologia* 1992; **35**: 1060-1067.
2. Kyvik KO, Green A, Beck-Nielsen H. Concordance for insulin-dependent diabetes mellitus: a population based study of young Danish twins. *Br Med J* 1995; **311**: 913-917.

3. Lorenzen T, Pociot F, Hougaard P, et al. Long term risk for IDDM in first-degree relatives of patients with IDDM. *Diabetologia* 1994; **37**: 321-327.
4. Risch N. Assessing the role of HLA-linked and unlinked determinants of disease. *Am J Hum Genet* 1987; **40**: 1-14.
5. Thorsby E, Rønningen KS. Particular HLA-DQ molecules play a dominant role in determining susceptibility or resistance to insulin-dependent diabetes mellitus. *Diabetologia* 1993; **36**: 371-377.
6. Rønningen KS. Genetics in the prediction of insulin-dependent diabetes mellitus: From theory to practice. *Ann Med* 1997; **29**: 387-392.
7. Undlien DE, Friede T, Rammensee H-G, et al. HLA-encoded genetic disposition in IDDM. DR4 subtypes may be associated with degrees of protection. *Diabetes* 1997; **46**: 143-149.
8. Rønningen KS, Keiding N, Green A. Correlations between the incidence of childhood-onset type I diabetes in Europe and HLA genotypes. *Diabetologia* 2001; **44**: B51-B59.
9. Bennett ST, Lucassen AM, Gough SCL, Powell EE, Undlien DE, Pritchard LE, Merriman ME, Kawaguchi Y, Dronsfield M, Pociot F, Nerup J, Bouzekri N, Cambon-Thomsen A, Rønningen KS, Barnett AH, Bain SC, Todd JA. Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nature Genet* 1995; **9**: 284-291.
10. Davies J, Kawaguchi Y, Bennett ST, et al. A genome wide search for human type 1 diabetes susceptibility genes. *Nature* 1994; **371**: 130-136.
11. Hashimoto L, Habita C, Beressi JP, et al. Genetic mapping of a susceptibility locus for insulin-dependent diabetes mellitus on chromosome 11q. *Nature* 1994; **371**: 161-164.
12. Field LL, Tobias R, Magnus T. A locus on chromosome 15q26 (IDDM 3) produces susceptibility to insulin-dependent diabetes mellitus. *Nature Genet* 1994; **8**: 189-194.
13. Copeman JB, Hearne CM, Cornall RJ, Reed PW, Cucca F, Rønningen KS, Undlien DE, Buzzetti R, Nistico L, Tosi R, Pociot F, Nerup J, Cornelis F, Barnett AH, Bain SC, Todd JA. Fine localisation of type 1 diabetes susceptibility gene (IDDM7) to human chromosome 2q31-q33 by linkage disequilibrium mapping. *Nature Genet* 1995; **9**: 80-85.
14. Reed PW, Cucca F, McKinney P, Bosi E, Jøner G, Rønningen KS, Thorsby E, Undlien DE, Barnett AH, Bain SC, Todd JA. Linkage disequilibrium with type 1 diabetes at human chromosome 10p11 localized on a 1 cM resolution microsatellite marker map of 10p11-10q11. *Hum Mol Genet* 1997; **6**: 1011-1016.
15. Merriman TR, Twells R, Merriman ME, Cox R, Cucca F, Jøner G, Rønningen KS, Thorsby E, Undlien DE, Bain S, Barnett AH, Todd JA. A strategy for identification of complex disease genes: localisation of type 1 diabetes polygene (IDDM6) to 1.7 cM of human chromosome 18q12. *Hum Mol Genet* 1997; **6**: 1003-1010.
16. Concannon P, Gogolin-Ewens KJ, Hind DA, et al. A second-generation screen of the human genome for susceptibility to insulin-dependent diabetes mellitus. *Nature Genet* 1998; **19**: 292-296.
17. European consortium for IDDM Genome Studies: A genomewide scan for type 1-diabetes susceptibility in Scandinavian families: identification of new loci with evidence for interaction. *Am J Hum Genet* 2001; **69**: 1301-1313.